REMARKS:

This is a full response to the outstanding nonfinal Office Action dated January 28, 2003. Applicant respectfully traverses and requests reconsideration.

All of the original 55 claims remain pending. Claims 1, 4, 8, 33, 34, and 40 were previously amended. There are no claims amended in the current Response. In the January 28, 2003 Office Action, claims 1-55 stand rejected as allegedly being obvious under 35 U.S.C. 103.

I. Summary of the Present Invention

Applicant Ira A. Gerson has invented a novel method and apparatus for processing an input speech signal during presentation of an output audio signal. At least one embodiment of the operation of Mr. Gerson's invention may be summarized as:

> A start of an input speech signal is detected during presentation of an output audio signal *and an input start time, relative to the output audio signal, is determined.* The input start time is *then provided for use in responding to the input speech signal.* In another embodiment, *the output audio signal has a corresponding identification.* When the input speech signal is detected during presentation of the output audio signal, *the identification of the output audio signal is provided* for use in responding to the input speech signal. Information signals comprising data and/or control signals are provided *in response to at least the contextual information provided, i.e., the input start time and/or the identification of the output audio signal.* In this manner, the present invention *accurately establishes a context* of an input speech signal relative to an output audio signal regardless of the delay characteristics of the underlying communication system.

Abstract of the current invention (emphasis added).

II. Amendment to the Specification

Applicant has made an amendment to paragraph [0052] to correct a typographical error. In the first line, "client-sever" has been amended to read "client-server." This change is to correct an editing error and does not add new matter to the specification.

III. Prior Art Made of Record

The prior art made of record has been considered, but is not believed to effect the patentability of the presently pending claims.

IV. Response to Rejections under 35 U.S.C. § 103

Claims 1-55 were rejected in the January 28, 2003 Office Action under 35 U.S.C. §103(a) as purportedly being obvious over U.S. Patent No. 5,765,130 issued to Nguyen (the *Nguyen* reference) in view of U.S. Patent No. 5,708,704 issued to Fisher (the *Fisher* reference). Applicant respectfully transverses the Examiner's rejection and requests reconsideration of the patentability of the invention.

According to MPEP § 706.02(j), for a claim to be obvious, there must be a) a suggestion or motivation to combine reference teachings, b) a reasonable expectation of success, and c) the references must teach all of the claim limitations, *In re Vaeck*, 947 F.2d 488, 20 U.S.P.Q.2d 1438 (Fed. Cir. 1991). It is well established at law that, for a proper rejection of a claim under 35 U.S.C. §103 as being obvious based upon a combination of references, the cited combination of references must disclose, teach, or suggest, either implicitly or explicitly, all elements/features/steps of the claim at issue. See, *e.g., In re Dow Chemical*, 5 U.S.P.Q.2d 1529, 1531 (Fed. Cir. 1988), and *In re Keller*, 208 U.S.P.Q.2d 871, 881 (C.C.P.A. 1981).

a. The *Nguyen* Reference

The *Nguyen* reference states that it provides:

> A barge-in detector for use in connection with a speech recognition system forms a prompt replica *for use in detecting the presence or absence of user input to the system*. The *replica is indicative of the prompt energy applied* to an input of the system. The detector detects the application of user input to the system, even if concurrent with a prompt, and enables the system to quickly respond to the user input.

Abstract of the *Nguyen* reference (emphasis added).

In the January 28, 2003 Office Action, the Examiner acknowledges that the *Nguyen* reference does not specifically teach providing the input time to establish a context in responding to the input speech signal. The *Nguyen* reference also does not appear to disclose: referencing the detected start of the user input signal to the audio output signal; determining, relative to the output audio signal, an input start time; establishing a context; and providing the input start time to establish a context in responding to the input speech signal. Thus, the *Nguyen* reference suffers from the defect found in prior art barge-in detectors of not taking into account the uncertain delay characteristics found in wireless and packet data systems (Specification, p. 2,

lines 2-5 and 21-28). The delays make it difficult to determine the portion of the output audio signal to which a user input signal corresponds.

b. The *Fisher* Reference

The *Fisher* reference states that it provides:

> A system for a voice prompt talkover capability. The system and method include a speech recognizer (36), an echo canceler (24), a delay finder (27) and a utterance extraction mechanism (30). The delay finder (27) determines the propagation delay between outgoing signal and the incoming signal and stores the signal in a delay buffer (23) in accordance with the delay and the filter length of the echo canceler. The echo canceler (24), using the incoming signals, the delay buffer and the delay enhance the incoming signal *by removing the echo*. The utterance extractor (30) determines the start of an utterance based on energy in the vectors of the LPC buffer (28). Upon the detection of an utterance the utterance extractor (30) initiates the start of recognition and terminates operation of the echo canceler (24) and the prompt play out (14). A recognition result is returned from the speech recognition processor (36) to the application (12). (emphasis added)

Abstract of the *Fisher* reference (emphasis added).

The Examiner states that *Fisher* teaches establishing input start time to establish a context in responding to the input speech signal. However, in contrast to the Examiner's assertion, The *Fisher* reference does not disclose: referencing the detected start of the user input signal to the audio output signal; determining, relative to the output audio signal, an input start time; establishing a context; and providing the input start time to establish a context in responding to the input speech signal.

c. In Regard to Independent Claim 1, the Combination of the *Nguyen* and *Fisher* References Fails to Teach all of the Applicant's Claim Limitations

Applicant's independent claim 1 identifies an embodiment of his invention as:

> A method for processing an input speech signal during presentation of an output audio signal, the method comprising steps of:
>
> detecting a start of the input speech signal;
>
> *determining, relative to the output audio signal, an input start time* of the start of the input speech signal; and
>
> providing the input start time *to establish a context in responding to the input speech signal.* (emphasis added)

Independent claim 1 is allowable for at least the reason that the combination of the *Nguyen* reference in view of the *Fisher* reference does not disclose, teach, or suggest the features that are highlighted in claim 1 above. More specifically, the references cited by the Examiner do not

disclose: (1) determining, relative to the output audio signal, an input start time; (2) establishing a context; and (3) providing the input start time to establish a context in responding to the input speech signal.

In rejecting claim 1, the Examiner states:

> As per claim 1, *Nguyen* teaches a method for processing an input speech signal during presentation of an output audio signal, the method comprising the steps of:
>
> detecting a start of the input speech signal; detecting, relative to the output audio signal, an input start time of the input speech signal; and, providing the input start time for use in responding to the input speech signal.

The Examiner cites to the Abstract of *Nguyen* to support the assertion noted immediately above. However, a review of the abstract of *Nguyen* (provided above) clearly shows that the *Nguyen* Abstract does not mention "detecting a start of the input speech signal," "input start time," or "providing the input start time for use in responding to the input speech signal."

The *Nguyen* reference describes an "apparatus for detecting the onset of user speech on a telephone line which also carries voice prompts for the user." Col. 1, lns. 11, 12. The *Nguyen* reference states, "Separating the user's voice response from the prompt is therefore a difficult task which has hitherto not been well handled." Col. 2, lns. 14-16. The *Nguyen* reference states that it "remove[s] the effects of the prompt residue from the line of a telephone system by predicting or modeling the time varying energy of the expected residue during successive sampling frames (occupying defined time intervals) over which the signal occurs and then subtracting that residue energy from the line input signal." Col. 2, lns. 35-41. The *Nguyen* reference is thus a filtering technique for more accurately detecting the presence of user speech.

Detecting the presence of user speech is a predicate for Applicant's invention. For this reason, Applicant cites the *Nguyen* reference as one means of allowing "a speech recognition system to respond to input speech while output speech is simultaneously being generated by the system." Paragraph [0040]. However, Applicant's invention includes *determining, relative to the output audio signal, an input start time* of the start of the input speech signal; and providing the input start time *to establish a context* in *responding to the input speech signal.* The *Nguyen* reference provides no disclosure or teachings regarding creating new information, *i.e.,* establishing a context, or what to do with the new information after the creation of the new information. One example of the difference between Applicant's invention and the *Nguyen*

reference is that Applicant's invention allows for using the contextual information to establish a context that can be used to determine whether the input speech signal was in response to the output audio signal used to determine the interval. The *Nguyen* reference fails to provide any teachings that could be used to correlate a particular output audio signal, *e.g.* a frame of data, with a prompt.

In the January 28, 2003 Office Action, the Examiner acknowledges that the *Nguyen* reference does not specifically teach providing the input time to establish a context in responding to the input speech signal. However, the Examiner states that the *Fisher* reference "does teach establishing input start time to establish a context in responding to the input speech signal." The Examiner cites the Abstract, FIG. 3B, and col. 3, ln. 59 through col. 4, ln. 64, of the *Fisher* reference to support his assertion regarding the teachings of the *Fisher* reference.

A review of the Abstract of the *Fisher* reference (provided above) clearly shows that the Abstract of the *Fisher* reference does not mention *establishing a context in responding to the input speech signal.* Similarly, the 72 line block cited by the Examiner and FIG. 3B also do not mention *establishing a context in responding to the input speech signal.* The *Fisher* reference is directed to a speech recognition method and system with voice-activated interrupt capability. The *Fisher* reference provides an echo canceler that cancels the echo of the outgoing prompt message from a portion of the incoming signals determined in accordance with the channel delay and a filter length associated with the echo canceler. See *Fisher*, Summary of the Invention. Similarly to the *Nguyen* reference, the *Fisher* reference is thus a filtering technique for more accurately detecting the presence of user speech.

Detecting the presence of user speech is a predicate for Applicant's invention. For this reason, Applicant cites the *Fisher* reference as one means of allowing "a speech recognition system to respond to input speech while output speech is simultaneously being generated by the system." Paragraph [0040]. However, Applicant's invention includes *determining, relative to the output audio signal, an input start time* of the start of the input speech signal; and providing the input start time *to establish a context* in *responding to the input speech signal.* The *Fisher* provides no disclosure or teachings regarding creating new information, *i.e.* establishing a context, or what to do with the new information after the creation of the new information. One example of the difference between Applicant's invention and the *Fisher* reference is that Applicant's invention allows for using the contextual information to establish a context that can

be used to determine whether the input speech signal was in response to the output audio signal used to determine the interval. *Fisher* reference fails to provide any teachings that could be used to correlate a particular output audio signal, *e.g.* a frame of data, with a prompt. The *Fisher* reference, within the portion cited by the Examiner, states, "The beginning and ending frames of utterances within the stored signals is determined using the utterance finding module 146. If utterances are detected in the stored signals, the present invention stops the playing of the outgoing prompt message and turns echo cancellation off at block 152." The *Fisher* reference does not teach correlating a particular output audio signal, *e.g.* a frame of data, with a prompt.

In addition to the combined references failing to teach all of the Applicant's claim limitations, the references fail to suggest or motivate one skilled in the art to combine the teachings of the references and fail to provide a reasonable expectation of success through a combination. The *Nguyen* reference and the *Fisher* reference are filtering technique for more accurately detecting the presence of user speech. There is no motivation to combine the references to address the problems addressed through Applicant's invention, *e.g.* providing the input start time to establish a context *in responding to* the input speech signal.

In the January 28, 2003 Office Action, the Examiner states, "It would have been obvious to one with ordinary skill in the art at the time of invention to establish a context in responding to the input speech signal as taught by *Fisher* into the method of *Nguyen*, because this would provide effective barge-in capability independent of the structure of the communications network configuration." However, neither *Nguyen* nor *Fisher* address establishing a context in responding to the input speech signal. Consequently, the combination of the *Nguyen* reference in view of the *Fisher* reference does not render claim 1 obvious, and the rejection should be withdrawn.

d. In Regard to Claim 2, the Combination of the *Nguyen* and *Fisher* References Fails to Teach all of the Applicant's Claim Limitations

Applicant's dependent claim 2 identifies an embodiment of his invention as:

> *The method of claim 1*, wherein the input start time comprises any one of a *time stamp relative to a temporal context of the output audio signal*, a *sample index relative to a sample context of the output audio signal*, and a *frame index relative to a frame context of the output audio signal*. (emphasis added)

Dependent claim 2 is allowable for at least the reason that the combination of the *Nguyen* reference in view of the *Fisher* reference does not disclose, teach, or suggest the features that are

highlighted in claim 2 above. More specifically, the references cited by the Examiner do not disclose, at the very least: (1) a time stamp relative to a temporal context of the output audio signal, (2) a sample index relative to a sample context of the output audio signal; and (3) a frame index relative to a frame context of the output audio signal.

In rejecting Claim 2, the Examiner states:

> As per claim 2, *Nguyen* teaches the method of claim 1, wherein the input start time comprises any of a time stamp relative to a temporal context of the output audio signal, a sample index relative to a sample context of the output audio signal, and a frame index relative to a frame context of the output audio signal.

The Examiner cites to col. 1, lns. 54-67, Col. 2, ln. 35 - col. 3, ln. 53. The portion of the *Nguyen* reference cited by the Examiner refers to filtering techniques used in identifying user input. The portions cited by the Examiner do not disclose or teach an input start time having any contextual relationship to the output audio signal. The *Nguyen* reference does not teach the use of time stamps, sample indexes, or frame indexes for any purpose. Consequently, the combination of the *Nguyen* reference in view of the *Fisher* reference does not render claim 2 obvious, and the rejection should be withdrawn.

e. <u>In Regard to Claim 4 and 5, the Combination of the *Nguyen* and *Fisher* References Fails to Teach Determining an Identification Corresponding to the Output Audio Signal</u>

Applicant's dependent claim 4 identifies an embodiment of his invention as:

> A method for processing an input speech signal during presentation of an output audio signal, the method comprising steps of:
>
> detecting the input speech signal;
>
> *determining an identification corresponding to the output audio signal*; and
>
> *providing the identification to establish a context in responding to the input speech signal.* (emphasis added)

Applicant's dependent claim 5 identifies an embodiment of his invention as:

> A computer-readable medium having computer-executable instructions for performing the steps recited in claim 4.

Dependent claim 4 is allowable for at least the reason that the combination of the *Nguyen* reference in view of the *Fisher* reference does not disclose, teach, or suggest the features that are highlighted in claim 4 above. More specifically, the references cited by the Examiner do not disclose, at the very least: (1) determining an identification corresponding to the output audio

signal; and (2) providing the identification to establish a context in responding to the input speech signal.

In rejecting claims 4 and 5, the Examiner states:

> Claims 4-5 are method claims similar in scope and content of method claims 1-3 and are rejected under similar rationale.

The Examiner fails to cite to any portion of either the *Nguyen* reference or the *Fisher* reference that teaches or discloses the step of determining an identification corresponding to the output audio signal. Claims 4-5 are not similar in scope and content to method claims 1-3. Claims 1-3 do not include the limitation of determining an identification corresponding to the output audio signal. Consequently, the combination of the *Nguyen* reference in view of the *Fisher* reference does not render claims 4 and 5 obvious, and the rejection should be withdrawn.

f. In Regard to Claim 2-5, the Combination of the *Nguyen* and *Fisher* References Fails to Teach all of the Applicant's Claim Limitations and these Dependent Claims are Patentable as they Depend from allowable Claims.

If independent claim 1 is allowable over the prior art of record, then its dependent claims 2-5 are allowable as a matter of law, because these dependent claims contain all features, elements, or steps of independent claim 1. *In re Fine*, 837 F.2d 1071 (Fed. Cir. 1988).

g. In Regard to Independent Claim 6, the Combination of the *Nguyen* and *Fisher* References Fails to Teach all of the Applicant's Claim Limitations

Applicant's independent claim 6 identifies an embodiment of his invention as:

> In a subscriber unit in wireless communication with an infrastructure comprising a speech recognition server, the subscriber unit comprising a speaker and a microphone, wherein the speaker provides an output audio signal and the microphone provides an input speech signal, a method for processing the input speech signal, the method comprising steps of:
>
> detecting a start of the input speech signal during presentation of the output speech signal;
>
> *determining, relative to the output audio signal, an input start time of the start of the input speech signal*; and
>
> *providing the input start time to the speech recognition server as a control parameter*. (emphasis added)

Independent claim 6 is allowable for at least the reason that the combination of the *Nguyen* reference in view of the *Fisher* reference does not disclose, teach, or suggest the features that are highlighted in claim 6 above. More specifically, the references cited by the Examiner do not

disclose, at the very least: (1) determining, relative to the output audio signal, an input start time of the start of the input speech signal; and (2) providing the input start time to the speech recognition server as a control parameter.

In rejecting claim 6, the Examiner states:

> As per claim 6, *Nguyen* teaches a method for processing input speech signal, comprising the steps of: detecting a start of the input speech signal during presentation of the output signal; determining, relative to the output audio signal, an input start time of the start of the input speech signal; and, providing the input start time to the speech server as control parameter.

The Examiner cites to the Abstract and col. 2, ln. 35 through col. 3, ln. 53 of the *Nguyen* reference to support the assertion noted immediately above. However, a review of the Abstract of the *Nguyen* reference (provided above) clearly shows that the Abstract and col. 2, ln. 35 through col. 3, ln. 53 do not mention "detecting a start of the input speech signal," "input start time," or "providing the input start time for use in responding to the input speech signal." Examiner's conclusions, unsupported by any correlation between the claim language and the references, cannot support the Examiner's rejection of claim 6. Consequently, the combination of the *Nguyen* reference in view of the *Fisher* reference does not render claim 6 obvious, and the rejection should be withdrawn.

h. <u>In Regard to Claim 7, the Combination of the *Nguyen* and *Fisher* References Fails to Teach all of the Applicant's Claim Limitations</u>

Applicant's dependent claim 7 identifies an embodiment of his invention as:

> The method of claim 6, further comprising a step of: *receiving at least one information signal from the speech recognition server based at least in part upon the input start time.* (emphasis added)

Dependent claim 7 is allowable for at least the reason that the combination of the *Nguyen* reference in view of the *Fisher* reference does not disclose, teach, or suggest the features that are highlighted in claim 7 above. More specifically, the references cited by the Examiner do not disclose, at the very least, receiving at least one information signal from the speech recognition server based at least in part upon the input start time.

In rejecting claims 7-12, the Examiner has simply listed the elements of claims 7-12 and concluded that the *Nguyen* reference teaches these elements. The Examiner cites to the Abstract, col. 2, ln. 35 through col. 3, ln.53, and col. 6, ln. 31 through col. 7, ln.17. Despite the length of the portions cited by the Examiner, the *Nguyen* reference fails to teach the elements of claims 7-

12. In relationship to claim 7, the *Nguyen* reference cited by the Examiner does not teach receiving at least one information signal from the speech recognition server based at least in part upon the input start time. The *Nguyen* reference does not teach calculating an input start time referenced to the output audio signal. Consequently, the combination of the *Nguyen* reference in view of the *Fisher* reference does not render claim 7 obvious, and the rejection should be withdrawn.

i. In Regard to Claim 8, the Combination of the *Nguyen* and *Fisher* References Fails to Teach all of the Applicant's Claim Limitations

Applicant's dependent claim 8 identifies an embodiment of his invention as:

> The method of claim 6, the step of determining the input start time further comprising the steps of:
>
> determining the input start time *no earlier than a start of the output audio signal and no later than a start of a subsequent output audio signal.* (emphasis added)

Dependent claim 8 is allowable for at least the reason that the combination of the *Nguyen* reference in view of the *Fisher* reference does not disclose, teach, or suggest the features that are highlighted in claim 8 above. More specifically, the references cited by the Examiner do not disclose, at the very least, determining the input start time no earlier than a start of the output audio signal and no later than a start of a subsequent output audio signal.

In rejecting claims 7-12, the Examiner has simply listed the elements of claims 7-12 and concluded that the *Nguyen* reference teaches these elements. In relationship to claim 8, the *Nguyen* reference cited by the Examiner does not teach determining the input start time no earlier than a start of the output audio signal and no later than a start of a subsequent output audio signal. The *Nguyen* reference does not teach calculating an input start time referenced to the output audio signal. Consequently, the combination of the *Nguyen* reference in view of the *Fisher* reference does not render claim 8 obvious, and the rejection should be withdrawn.

j. In Regard to Claim 7-12, the Combination of the *Nguyen* and *Fisher* References Fails to Teach all of the Applicant's Claim Limitations and these Dependent Claims are Patentable as they Depend from allowable Claims.

If independent claim 6 is allowable over the prior art of record, then its dependent claims 7-12 are allowable as a matter of law, because these dependent claims contain all features, elements, or steps of independent claim 6. *In re Fine*, 837 F.2d 1071 (Fed. Cir. 1988).

k. In Regard to Claim 13 through 55, the Combination of the *Nguyen* and *Fisher* References Fails to Teach all of the Applicant's Claim Limitations and these Dependent Claims are Patentable as they Depend from allowable Claims.

The Examiner states that claims 13 through 17 are method claims similar in scope and content of method claims 6 through 12. The Examiner rejected claims 13 through 17 under similar rational to that provided in regard to claims 6 through 12

The Examiner states that claims 18 through 30 are method claims similar in scope and content of method claims 6 through 12. The Examiner rejected claims 18 through 30 under similar rational to that provided in regard to claims 6 through 12

The Examiner states that claims 31 through 55 are apparatus claims that implement the method claims 6 through 30. The Examiner further states that claims 31 through 55 are similar in scope and content of method claims 6 through 30. The Examiner rejected claims 31 through 55 under similar rational to that provided in regard to claims 6 through 30.

Applicant respectfully disagrees with the Examiner's characterizations of Applicant's 42 claims numbered 13 through 55. The 42 claims cannot be so generally characterized. For example, in claim 6 an input speech signal is detected within a subscriber unit during the presentation of an output speech signal. In claim 17, an information signal based at least in part on an identification and a parameterized speech signal is received by a subscriber unit. In claim 18, an input start time is received from a subscriber unit by a speech recognition server. In claim 21, information signals are directed to a subscriber unit to control the operation of the subscriber unit. In claims 22 and 28, a subscriber unit is coupled to at least one device and information signals are directed to the at least one device to control operation of the at least one device. These features are not identical to the features in claims 1 through 12. These features illustrate some of the additional features of claims 31-50 that are not taught or suggested by the cited references. Consequently, for these reasons, as well as those discussed in connection with claims 1-12, *supra*, the combination of the *Nguyen* reference in view of the *Fisher* reference does not render claims 13 through 55 obvious, and the rejection should be withdrawn.

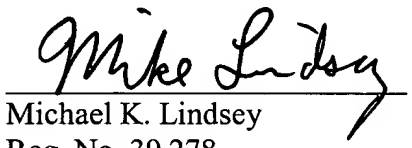
V. Conclusion

In light of the forgoing remarks, and for at least the reasons set forth above, Applicant respectfully submits that all rejections have been traversed or rendered moot and that now pending claims 1-55 are in condition for allowance. In summary, because neither the *Nguyen*

reference, nor the *Fisher* reference, suggests combining the filtering techniques described in those references to address the problems addressed by Applicant's invention, and because the two references do not teach all of the claim limitations for pending claims 1-55, those claims are in condition for allowance. Accordingly, Applicant requests that the rejections of claims 1-55 under 35 U.S.C. § 103 be withdrawn. If, in the opinion of the Examiner, a telephonic conference would expedite the examination of this matter, the Examiner is invited to call the undersigned attorney at (312) 595-1239.

Respectfully submitted,

Dated: October 22, 2003



Michael K. Lindsey
Reg. No. 39,278
Customer No. 22930
ATTORNEY FOR ASSIGNEE,
FastMobile, Inc.

Howrey Simon Arnold & White, LLP
321 North Clark Street, Suite 3400
Chicago, Illinois 60610
(312) 595-1239 (Tel)
(312) 595-2250 (Fax)